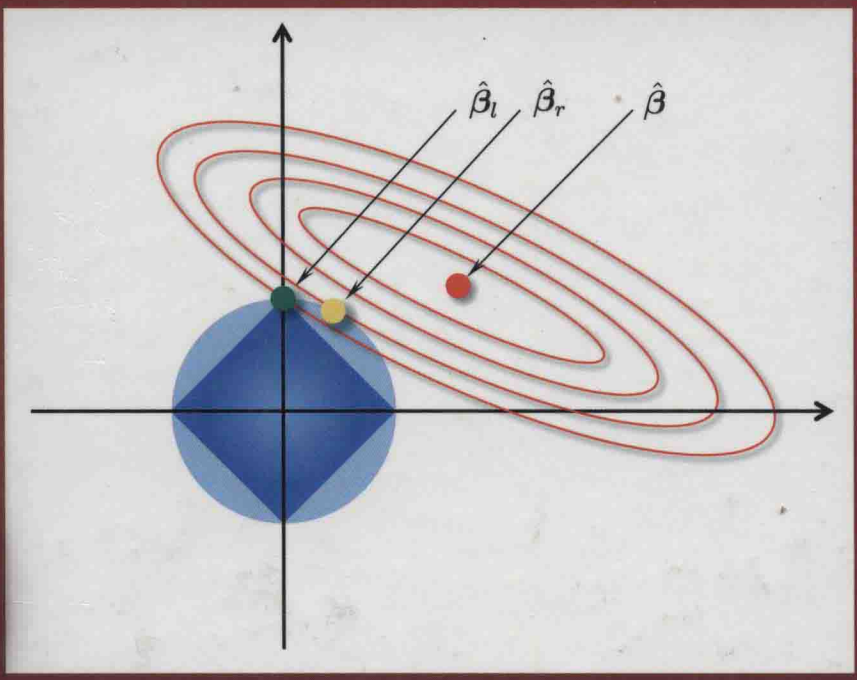
Texts in Statistical Science

Statistical Inference

An Integrated Approach

Second Edition



Helio S. Migon
Dani Gamerman
Francisco Louzada

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Texts in Statistical Science

Statistical Inference

An Integrated Approach

Second Edition

常州大学山书馆
藏书章

Helio S. Migon
Dani Gamerman
Francisco Louzada

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2015 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20140520

International Standard Book Number-13: 978-1-4398-7880-4 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Migon, Helio dos Santos.
Statistical inference : an integrated approach / Helio S. Migon, Dani Gamerman, Francisco Louzada-Neto. -- Second edition.
pages cm
"A CRC title."
Includes bibliographical references and index.
ISBN 978-1-4398-7880-4 (hardcover : alk. paper) 1. Mathematical statistics. I. Gamerman, Dani. II. Louzada-Neto, Francisco. III. Title.

QA276.M442 2014
519.5--dc232014019200 2014019200

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Statistical Inference

An Integrated Approach

Second Edition

CHAPMAN & HALL/CRC
Texts in Statistical Science Series

Series Editors
Francesca Dominici, *Harvard School of Public Health, USA*
Julian J. Faraway, *University of Bath, UK*
Martin Tanner, *Northwestern University, USA*
Jim Zidek, *University of British Columbia, Canada*

Statistical Theory: A Concise Introduction
F. Abramovich and Y. Ritov

Practical Multivariate Analysis, Fifth Edition
A. Afifi, S. May, and V.A. Clark

Practical Statistics for Medical Research
D.G. Altman

Interpreting Data: A First Course in Statistics
A.J.B. Anderson

Introduction to Probability with R
K. Baclawski

Linear Algebra and Matrix Analysis for Statistics
S. Banerjee and A. Roy

Statistical Methods for SPC and TQM
D. Bissell

Introduction to Probability
J. K. Blitzstein and J. Hwang

Bayesian Methods for Data Analysis, Third Edition
B.P. Carlin and T.A. Louis

Second Edition
R. Caulcutt

The Analysis of Time Series: An Introduction, Sixth Edition
C. Chatfield

Introduction to Multivariate Analysis
C. Chatfield and A.J. Collins

Problem Solving: A Statistician's Guide, Second Edition
C. Chatfield

Statistics for Technology: A Course in Applied Statistics, Third Edition
C. Chatfield

Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians
R. Christensen, W. Johnson, A. Branscum, and T.E. Hanson

Modelling Binary Data, Second Edition
D. Collett

Modelling Survival Data in Medical Research, Second Edition
D. Collett

Introduction to Statistical Methods for Clinical Trials
T.D. Cook and D.L. DeMets

Applied Statistics: Principles and Examples
D.R. Cox and E.J. Snell

Multivariate Survival Analysis and Competing Risks
M. Crowder

Statistical Analysis of Reliability Data
M.J. Crowder, A.C. Kimber, T.J. Sweeting, and R.L. Smith

An Introduction to Generalized Linear Models, Third Edition
A.J. Dobson and A.G. Barnett

Nonlinear Time Series: Theory, Methods, and Applications with R Examples
R. Douc, E. Moulines, and D.S. Stoffer

Introduction to Optimization Methods and Their Applications in Statistics
B.S. Everitt

Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models
J.J. Faraway

Linear Models with R, Second Edition
J.J. Faraway

A Course in Large Sample Theory
T.S. Ferguson

Multivariate Statistics: A Practical Approach
B. Flury and H. Riedwyl

Readings in Decision Analysis
S. French

Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Second Edition
D. Gamerman and H.F. Lopes

此为试读，需要完整PDF请访问：www.ertongbook.com

Bayesian Data Analysis, Third Edition
A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin

Multivariate Analysis of Variance and Repeated Measures: A Practical Approach for Behavioural Scientists
D.J. Hand and C.C. Taylor

Practical Data Analysis for Designed Practical Longitudinal Data Analysis
D.J. Hand and M. Crowder

Logistic Regression Models
J.M. Hilbe

Richly Parameterized Linear Models: Additive, Time Series, and Spatial Models Using Random Effects
J.S. Hodges

Statistics for Epidemiology
N.P. Jewell

Stochastic Processes: An Introduction, Second Edition
P.W. Jones and P. Smith

The Theory of Linear Models
B. Jørgensen

Principles of Uncertainty
J.B. Kadane

Graphics for Statistics and Data Analysis with R
K.J. Keen

Mathematical Statistics
K. Knight

Introduction to Multivariate Analysis: Linear and Nonlinear Modeling
S. Konishi

Nonparametric Methods in Statistics with SAS Applications
O. Korosteleva

Modeling and Analysis of Stochastic Systems, Second Edition
V.G. Kulkarni

Exercises and Solutions in Biostatistical Theory
L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Exercises and Solutions in Statistical Theory
L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Design and Analysis of Experiments with SAS
J. Lawson

A Course in Categorical Data Analysis
T. Leonard

Statistics for Accountants
S. Letchford

Introduction to the Theory of Statistical Inference
H. Liero and S. Zwanzig

Statistical Theory, Fourth Edition
B.W. Lindgren

Stationary Stochastic Processes: Theory and Applications
G. Lindgren

The BUGS Book: A Practical Introduction to Bayesian Analysis
D. Lunn, C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter

Introduction to General and Generalized Linear Models
H. Madsen and P. Thyregod

Time Series Analysis
H. Madsen

Pólya Urn Models
H. Mahmoud

Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition
B.F.J. Manly

Introduction to Randomized Controlled Clinical Trials, Second Edition
J.N.S. Matthews

Statistical Methods in Agriculture and Experimental Biology, Second Edition
R. Mead, R.N. Curnow, and A.M. Hasted

Statistics in Engineering: A Practical Approach
A.V. Metcalfe

Statistical Inference: An Integrated Approach, Second Edition
H. S. Migon, D. Gamerman, and F. Louzada

Beyond ANOVA: Basics of Applied Statistics
R.G. Miller, Jr.

A Primer on Linear Models
J.F. Monahan

Applied Stochastic Modelling, Second Edition
B.J.T. Morgan

Elements of Simulation
B.J.T. Morgan

Probability: Methods and Measurement
A. O'Hagan

Introduction to Statistical Limit Theory
A.M. Polansky

Applied Bayesian Forecasting and Time Series Analysis
A. Pole, M. West, and J. Harrison

Statistics in Research and Development, Time Series: Modeling, Computation, and Inference
R. Prado and M. West

Introduction to Statistical Process Control
P. Qiu

Sampling Methodologies with Applications
P.S.R.S. Rao

A First Course in Linear Model Theory
N. Ravishanker and D.K. Dey

Essential Statistics, Fourth Edition
D.A.G. Rees

Stochastic Modeling and Mathematical Statistics: A Text for Statisticians and Quantitative
F.J. Samaniego

Statistical Methods for Spatial Data Analysis
O. Schabenberger and C.A. Gotway

Bayesian Networks: With Examples in R
M. Scutari and J.-B. Denis

Large Sample Methods in Statistics
P.K. Sen and J. da Motta Singer

Decision Analysis: A Bayesian Approach
J.Q. Smith

Analysis of Failure and Survival Data
P. J. Smith

Applied Statistics: Handbook of GENSTAT Analyses
E.J. Snell and H. Simpson

Applied Nonparametric Statistical Methods, Fourth Edition
P. Sprent and N.C. Smeeton

Data Driven Statistical Methods
P. Sprent

Generalized Linear Mixed Models: Modern Concepts, Methods and Applications
W. W. Stroup

Survival Analysis Using S: Analysis of Time-to-Event Data
M. Tableman and J.S. Kim

Applied Categorical and Count Data Analysis
W. Tang, H. He, and X.M. Tu

Elementary Applications of Probability Theory, Second Edition
H.C. Tuckwell

Introduction to Statistical Inference and Its Applications with R
M.W. Trosset

Understanding Advanced Statistical Methods
P.H. Westfall and K.S.S. Henning

Statistical Process Control: Theory and Practice, Third Edition
G.B. Wetherill and D.W. Brown

Generalized Additive Models: An Introduction with R
S. Wood

Epidemiology: Study Design and Data Analysis, Third Edition
M. Woodward

Experiments
B.S. Yandell

To Mirna, our sons and grandchildren (H.S.M.)
To Bernardo (i.m.) and Violeta (D.G.)
To Gleici, Lucas, Caio, Samuel and Francisco (F.L.)

Preface to the Second Edition

It has been more than a decade since the 1st edition was released. Over this period the need for a more balanced account of the main schools of thought to statistical inference became clear. The main purpose of the book has always been a balanced and integrated presentation of these approaches to the subject. But the first edition was written only by the first two authors of this edition; both were assumedly Bayesian and mentioned that in the Preface for that edition.

Experience gathered in the meantime indicated to us some ingredients to get back to our original goal. Certainly a more detailed and comprehensive presentation of the pros and cons of each approach had to be pursued. The inclusion of a co-author who was more involved with the frequentist school was also recommended to achieve a more balanced presentation. We were fortunate to add both ingredients and present the current version of our work after these additions.

We have tried to eliminate (or reduce, more realistically) the typos present in the 1st edition. Additionally, we have included more explanation, more illustration and more exercises for the concepts already present here. More importantly, new material has been included. The main novelties are associated with a new section on empirical Bayes and penalized likelihoods and their impact on the regression models in the last chapter. Interestingly, these are topics that lie closer to the border between the two schools of thought. The content of this new material has also been connected to many other parts of the book, throughout its chapters. Additionally, we have expanded material in many sections, specially on hypothesis testing, method of moments and bias correction.

This has been an interesting journey for the authors that is far from over. Statistics is alive and growing as a discipline with an ever increasing impact of the growth in computational power. We have no illusion of having reached a complete and final text on the subject, but hope to have reached a stage where readers can grasp the main task ahead of them in their search for an integrated understanding of Statistical Inference.

There are a number of people that we would like to thank and show our gratitude. We all thank our common friend Luis Raul Pericchi for the suggestion of the 3rd author name and many other suggestions. Thanks are also due to our lifelong friend Basilio Pereira for much and relentless advice with references and perspective. There is no way to avoid acknowledging the fundamental help of Rob Calver. He has been keen on making this project viable since the release of the 1st edition. There were many many moments that he was the only one believing it. His persistence and enthusiasm was crucial to have us reached this stage. He is complemented by a number of competent colleagues at CRC who have helped us in many different ways. We also thank colleagues and students who have helped shape the work in various ways, and our families for their encouragement and support. To all of them, our sincerest gratitude. We hold none of them responsible and take full account for the views expressed in this book.

H.S.M., D.G. and F.L.
Rio de Janeiro and São Carlos, May 2014

此为试读,需要完整PDF请访问: www.ertongbook.com

Preface to the First Edition

This book originated from the lecture notes of a course in Statistical Inference taught at the M.Sc. programs in Statistics at UFRJ and IMPA (once). These have been used since 1987. During this period, various modifications have been introduced until we arrived at this version, judged as minimally presentable.

The motivation to prepare this book came from two different sources. The first and more obvious one for us was the lack of texts in Portuguese, dealing with statistical inference to the desired depth. This motivation led us to prepare the first draft of this book in the Portuguese language in 1993. The second, and perhaps the most attractive as a personal challenge, was the perspective adopted in this text. Although there are various good books in the literature dealing with this subject, in none of them could we find an integrated presentation of the two main schools of statistical thought: the frequentist (or classical) and the Bayesian. This second motivation led to the preparation of this English version. This version has substantial changes with respect to the Portuguese version of 1993. The most notable one is the inclusion of a whole new chapter dealing with approximation and computationally intensive methods.

Generally, statistical books follow their author's point of view, presenting at most, and in separate sections, related results from the alternative approaches. In this book, our proposal was to show, wherever possible, the parallels existing between the results given by both methodologies. *Comparative Statistical Inference* by V. D. Barnett (1973) is the book that is closest to this proposal. It does not, however, present many of the basic inference results that should be included in a text proposing a wide study of the subject. Also we wanted to be as comprehensive as possible for our aim of writing a textbook in statistical inference.

This book is organized as follows. The first chapter is an introduction, presenting the readers with the way we find most appropriate to think of Statistics: discussing the concept of information. Chapter 2 presents some basic concepts of statistics such as sufficiency, exponential family, Fisher information, permutability and likelihood functions. Another basic concept specific

to Bayesian inference is prior distribution; that is separately dealt with in Chapter 3.

Certain aspects of inference are individually presented in Chapters 4, 6, and 7. Chapter 4 deals with parameter estimation where, intentionally, point and interval estimation are presented as responses to the summarization question, and not as two unrelated procedures. The important results for the Normal distribution are presented and also serve an illustrative purpose. Chapter 6 is about hypotheses testing problems under the frequentist approach and also under the various possible forms of the Bayesian paradigm.

In between them, lies Chapter 5 where all approximation and computationally based results are gathered. The reader will find there at least a short description of the main tools used to approximately solve the relevant statistical problem for situations where an explicit analytic solution is not available. For this reason, asymptotic theory is also included in this chapter.

Chapter 7 covers prediction from both the frequentist and Bayesian points of view, and includes the linear Bayes method. Finally in Chapter 8, an introduction to Normal linear models is made. Initially the frequentist approach is presented, followed by the Bayesian one. Based upon the latter approach, generalizations are presented leading to the hierarchical and dynamic models.

We would like to alert the readers from the onset that the our preferred point of view is the Bayesian one. So, a natural emphasis is given to this approach. We tried, however, to develop a critical analysis and to present the most important results of both approaches commenting on the positive and negative aspects of both. As it has already been said, the level of this book is adequate for an M.Sc. course in Statistics, although we do not rule out the possibility of its use in an advanced undergraduate course aiming to compare the two approaches.

This book can also be useful for the more mathematically trained professionals from related areas of Science such as Economics, Mathematics, Engineering, Operations Research and Epidemiology. The basic requirements are knowledge of calculus and probability, although basic notions of linear algebra are also used. As this book is intended as a basic text in statistical inference, various exercises are included at the end of each chapter. We have also included sketched solutions to some of the exercises and a list of distributions at the end of the book, for easy reference.

There are many possible uses of this book as a textbook. The first and most obvious one is to present all the material in the order it appears in the book and without skipping sections. This may be a heavy workload for a one-semester course. In this case we suggest dropping Chapter 8 for a later

course. A second option for exclusion in a first course is Chapter 5, although we strongly recommend it for anybody interested in the modern approach to statistics, geared towards applications. The book can also be used as a text for a course that is more strongly oriented towards one of the schools of thought. For a Bayesian route, follow Chapters 1, 2, 3, Sections 4.1, 4.4.1 and 4.5, Chapter 5, Sections 6.3, 6.4, 6.5, 7.1, 7.3.1, 7.4, 8.1, 8.3, 8.4 and 8.5. For a classical route, follow Chapter 1, Sections 2.1, 2.2, 2.5, 2.6, 4.2, 4.3, 4.4.2 and 4.5, Chapter 5, Sections 6.1, 6.2, 6.4, 6.5, 7.2, 7.3.2, 7.4, 8.1 and 8.2.

This book would not have been possible without the cooperation of various people. An initial and very important impulse was the typing of the original lecture notes in TeX by Ricardo Sandes Ehlers. Further help was provided by Ana Beatriz Soares Monteiro, Carolina Gomes, Eliane Amiune Camargo, Monica Magnanini and Otávio Santos Figueiredo. Besides these, many of our former students helped with suggestions and criticism. Careful proofreading of this manuscript was made by our past M.Sc. students and present colleagues Alexandra Mello Schmidt, Hedibert Freitas Lopes and Marco Antonio Rosa Ferreira. Many useful suggestions and comments were provided at this later stage by Steve Brooks, Eduardo Gutierrez-Peña and Gabriel Huerta. We also had the stimulus of several colleagues; in particular, we would like to mention Basílio de B. Pereira. I would also like to thank Nicki Dennis for her support and encouragement throughout all the stages of preparation of this book and for making us feel at home with Arnold. Our families also played the important roles of support and understanding, especially in the weekends and late nights spent trying to meet deadlines! To all of them, our gratitude.

Finally, the subject of the book is not new and we are not claiming any originality here. We would like to think that we are presenting the subject in a way that is not favored in many textbooks and that will help the readers to have an integrated view of the subject. In our path to achieve this goal, we have had the influence of many researchers and books. We have tried to acknowledge this influence by referring to these books whenever we felt it provided a worthwhile reading description of a topic. Therefore, for every major subject presented in our book we tried to relate it to books that treated the subject in a more complete or more interesting way. In line with a textbook character, we opted to favor books rather than research papers as references. We would like to think of our book as a basis for discovery and will feel our task is acomplished whenever readers understand the subject through the book alone, its references or a combination of both.

H.S.M. & D.G.
Rio de Janeiro, December 1998

Contents

Preface to the Second Edition xiii

Preface to the First Edition xv

1 Introduction 1
- 1.1 Information 2
- 1.2 The concept of probability 3
- 1.3 Assessing subjective probabilities 6
- 1.4 An example 7
- 1.5 Linear algebra and probability 12
  - 1.5.1 Probability theory 12
  - 1.5.2 Linear algebra 15
- 1.6 Notation 16
- 1.7 Outline of the book 17

Exercises 19

2 Elements of inference 21
- 2.1 Common statistical models 21
- 2.2 Likelihood-based functions 23
  - 2.2.1 Likelihood function 24
  - 2.2.2 Fisher information 27
- 2.3 Bayes theorem 31
  - 2.3.1 Prediction 32
  - 2.3.2 Sequential nature of the Bayes theorem 35
- 2.4 Exchangeability 37
- 2.5 Sufficiency and exponential family 41
- 2.6 Parameter elimination 47
  - 2.6.1 Bayesian approach 47
  - 2.6.2 Classical approach 51

Exercises 55

3 Prior distribution **65**
3.1 Entirely subjective specification 65
3.1.1 Aspects on elicitation of expert beliefs 68
3.2 Specification through functional forms 69
3.3 Conjugacy with the exponential family 74
3.3.1 Binomial distribution 75
3.3.2 Normal distribution with known variance 75
3.3.3 Poisson distribution 76
3.3.4 Exponential distribution 76
3.3.5 Multinomial distribution 76
3.3.6 Normal distribution with known mean and unknown variance 77
3.3.7 Normal distribution with unknown mean and variance 78
3.4 Non-informative priors 80
3.4.1 Jeffreys non-informative priors 81
3.4.2 Reference priors 87
3.5 Hierarchical priors 89
Exercises 91

4 Estimation **99**
4.1 Introduction to decision theory 99
4.1.1 Classical decision theory 101
4.1.2 Bayesian decision theory 102
4.1.3 Statistical inference as a decision problem 104
4.2 Bayesian point estimation 105
4.3 Classical point estimation 109
4.3.1 Maximum likelihood 109
4.3.2 Method of least squares 113
4.3.3 Method of moments 115
4.3.4 Empirical distribution function 118
4.4 Empirical Bayes estimation 119
4.5 Comparison of estimators 124
4.5.1 Bias 124
4.5.2 Risk 128
4.5.3 Consistency 133
4.6 Interval estimation 135
4.6.1 Bayesian approach 135
4.6.2 Classical approach 139
4.6.3 Bonferroni inequalities 142

此为试读,需要完整PDF请访问: www.ertongbook.com